

Robustness in Stochastic Programs with Risk Constraints

Jitka Dupačová and Miloš Kopa

Dept. of Probability and Mathematical Statistics,
Faculty of Mathematics and Physics
Charles University, Prague,
Czech Republic
www.karlin.mff.cuni.cz/~kopa

- The model – constraints dependent on probability distributions
- Example: Risk-shaping with CVaR
- Stability results and contamination bounds
- Numerical example

The model

We shall deal with robustness properties of risk constrained stochastic programs

$$\min_{x \in \mathcal{X}} F_0(x, P)$$

subject to

$$F_j(x, P) \leq 0, j = 1, \dots, J; \quad (1)$$

- $\mathcal{X} \subset \mathbb{R}^n$ is a fixed nonempty convex set,
- functions $F_j(x, P)$, $j = 0, \dots, J$ may depend on P
- P is the probability distribution of a random vector ω with range $\Omega \subset \mathbb{R}^m$

Denote $\mathcal{X}(P)$ set of feasible solutions, $\mathcal{X}^*(P)$ set of optimal solutions, $\varphi(P)$ optimal value of the objective function in (1).

Contrary to usual [basic model](#)

$$\min_{x \in \mathcal{X}} E_P f(x, \omega) \quad (2)$$

constraints depend on P & structure is easier than for chance constraints.

EXAMPLE – Risk-shaping with CVaR [R-U]

Denote $f(x, \omega)$ random loss defined on $\mathcal{X} \times \Omega$ and $\alpha \in (0, 1)$ selected confidence level. **Conditional Value at Risk** at the confidence level α , CVaR_α , is defined as the mean of α -tail distribution of $f(x, \omega)$ and according to [R-U] can be evaluated by minimization of auxiliary function

$$\Phi_\alpha(x, v, P) := v + \frac{1}{1 - \alpha} E_P(f(x, \omega) - v)^+$$

with respect to v . Function $\Phi_\alpha(x, v, P)$ is linear in P and convex in v . If $f(x, \omega)$ is a *convex* function of x , $\Phi_\alpha(x, v, P)$ is convex jointly in (v, x) . **Risk-shaping with CVaR** handles several probability thresholds $\alpha_1, \dots, \alpha_J$ and loss tolerances $b_j, j = 1, \dots, J$. The problem is minimize $F(x)$ subject to $x \in \mathcal{X}$ and $\text{CVaR}_{\alpha_j}(x, P) \leq b_j, j = 1, \dots, J$. According to Theorem 16 of [R-U], this is equivalent to

$$\min_{x, v_1, \dots, v_J} \{F(x) \mid x \in \mathcal{X}, \Phi_{\alpha_j}(x, v_j, P) \leq b_j, j = 1, \dots, J\},$$

i.e. problem of the form (1) with expectation type constraints. Another example is mean-CVaR optimization problem.

Stability results and robustness wrt. P

Complete knowledge of the probability distribution is rare in practice – stability, robustness, output analysis wrt. P is needed .

- **Quantitative stability** cf. Theorem 5 of **Römisch (2003)** applied to (1) provides upper semicontinuity of the set of optimal solutions and a local Lipschitz property of the optimal value function for stochastic programs (1) with smooth, convex objective and one expectation type smooth convex constraint $F(x, P) \leq 0$ if at the optimal solution $x^*(P)$ of the unperturbed problem

$$\min_{x \in \mathcal{X}} F_0(x, P) \text{ s.t. } F(x, P) := E_P f(x, \omega) \leq 0$$

the constraint is not active, or if $\nabla F(x^*(P), P) \neq 0$.

To get metric regularity for multiple expectation type smooth convex constraints $F_j(x, P) \leq 0, j = 1, \dots, J$, general constraint qualification should be used or constraints reformulated as

$F(x, P) := \max_j F_j(x, P) \leq 0$ – again convex function.

- **Contamination bounds for the optimal value function**
- Another possibility – incorporate the incomplete knowledge of P into the model – **ambiguity, minimax**.

Robustness analysis via Contamination

was derived for (2), i.e. for $\mathcal{X}(P)$ independent of P and for expectation type objective $F_0(x, P)$.

Assume that SP (2) was solved for P , denote $\varphi(P)$ optimal value.

Changes in probability distribution P are modeled using **contaminated distributions** P_t ,

$$P_t := (1 - t)P + tQ, t \in [0, 1]$$

with Q another *fixed* probability distribution.

Via contamination, robustness analysis wrt. changes in P gets reduced to much simpler analysis wrt. scalar parameter t (see e.g. **resist**).

Objective function in (2) is linear in $P \implies F_0(x, P_t)$ is linear wrt. $t \implies$ optimal value function

$$\varphi(t) := \min_{x \in \mathcal{X}} F_0(x, P_t)$$

is **concave** on $[0, 1] \implies$ continuity and existence of directional derivatives in $(0, 1)$. Continuity at $t = 0$ is property related with stability for SP (2).

In general, one needs set of optimal solutions $\mathcal{X}^*(P) \neq \emptyset$, bounded.

Contamination Bounds

Concave $\varphi(t) \implies$ **contamination bounds**

$$\varphi(0) + t\varphi'(0^+) \geq \varphi(t) \geq (1-t)\varphi(0) + t\varphi(1), \quad t \in [0, 1]. \quad (3)$$

Using **arbitrary** optimal solution $x(P)$ of (2) \rightarrow upper bound

$$\varphi'(0^+) \leq F(x(P), Q) - \varphi(0).$$

Contamination bounds (3) are global, valid for all $t \in [0, 1]$. They **quantify** the change in optimal value due to considered perturbations of (2); cf. application to stress test of CVaR. The approach can be generalized to objective functions $F(x, P)$ convex in x and **concave in P** .

Stress testing and robustness analysis via contamination with respect to changes in probability distribution P is straightforward for **expected disutility models** (objective function is linear in P). Also stress testing for convex risk or deviation measures via contamination can be developed: When the risk or deviation measures are **concave with respect to probability distribution P** they are concave wrt. parameter t of contaminated probability distributions P_t .

Contamination bounds – constraints dependent on P

New problems – $\varphi(t)$ is no more concave in t .

Use $P_t := (1 - t)P + tQ$, $t \in (0, 1)$ in SP (1) at the place of P . Set of feasible solutions of (1) for contaminated probability distribution P_t

$$\mathcal{X}(P_t) = \mathcal{X} \cap \{x \mid F_j(x, P_t) \leq 0, j = 1, \dots, J\}. \quad (4)$$

Denote $\mathcal{X}(t)$, $\varphi(t)$, $\mathcal{X}^*(t)$ the set of feasible solutions, the optimal value $\varphi(P_t)$ and the set of optimal solutions $\mathcal{X}^*(P_t)$ of contaminated problem

$$\text{minimize } F_0(x, P_t) \text{ on the set } \mathcal{X}(P_t). \quad (5)$$

The task is to construct computable lower and upper bounds for $\varphi(t)$ & exploit them for robustness analysis in risk-shaping with CVaR or for a stochastic dominance test with respect to inclusion of additional scenarios. Thanks to the assumed structure of perturbations

- lower bound can be derived for $F_j(x, P)$, $j = 0, \dots, J$, linear or concave with respect to P without any smoothness or convexity assumptions with respect to x ,
- convexity of SP (1) is essential for directional differentiability of the optimal value function,
- further assumptions are needed for derivation of the upper bound.

Lower bound

1. One constraint dependent on P and objective F_0 independent of P :

$$\min_{x \in \mathcal{X}} F_0(x) \text{ subject to } F(x, P) \leq 0. \quad (6)$$

For contaminated probability distribution P_t we get

$$\min_{x \in \mathcal{X}} F_0(x) \text{ subject to } F(x, t) := F(x, P_t) \leq 0 \quad (7)$$

– **nonlinear parametric program** with scalar parameter $t \in [0, 1]$, set of feasible solutions $\mathcal{X}(t) := \{x \in \mathcal{X} \mid F(x, t) \leq 0\}$ **depends on t** .

In general, the optimal value function is not concave.

Theorem

Let $F(x, \bullet)$ be concave function of $t \in [0, 1]$. Then the optimal value function of (7)

$$\varphi(t) := \min_{x \in \mathcal{X}} F_0(x) \text{ subject to } F(x, t) \leq 0$$

is quasiconcave in $t \in [0, 1]$ with the lower bound

$$\varphi(t) \geq \min\{\varphi(1), \varphi(0)\}. \quad (8)$$

Proof is based on inclusion

$$\mathcal{X}((1-\lambda)t_1+\lambda t_2) \subset \{x \in \mathcal{X} \mid (1-\lambda)F(x, t_1)+\lambda F(x, t_2) \leq 0\} \subset \mathcal{X}(t_1) \cup \mathcal{X}(t_2) \quad (9)$$

valid for arbitrary $t_1, t_2 \in [0, 1]$ and $0 \leq \lambda \leq 1$.

2. When also **objective function depends** on probability distribution, i.e. on contamination parameter t , the problem is

$$\min_{x \in \mathcal{X}} F_0(x, t) := F_0(x, P_t) \text{ subject to } F(x, t) \leq 0. \quad (10)$$

For $F_0(x, P)$ linear or concave in P , lower bound can be obtained by application of the above quasiconcavity result (8) separately to $F_0(x, P)$ and $F_0(x, Q)$:

$$\varphi(t) \geq (1-t) \min\{\varphi(0), \min_{\mathcal{X}(Q)} F_0(x, P)\} + t \min\{\varphi(1), \min_{\mathcal{X}(P)} F_0(x, Q)\}. \quad (11)$$

The bound is more complicated but still computable.

3. For **multiple constraints** and contaminated probability distribution it would be necessary to prove first the inclusion $\mathcal{X}(t) \subset \mathcal{X}(0) \cup \mathcal{X}(1)$ and then the lower bound (8) for the optimal value $\varphi(t) = \min_{x \in \mathcal{X}(t)} F_0(x, P_t)$ can be obtained as in the case of one constraint.

Denote $\mathcal{X}_j(t) = \{x \mid F_j(x, P_t) \leq 0\}$. Then according to (9), $\mathcal{X}_j(t) \subset \mathcal{X}_j(0) \cup \mathcal{X}_j(1)$, hence

$$\mathcal{X}(t) \subset \mathcal{X} \cap \bigcap_j [\mathcal{X}_j(P) \cup \mathcal{X}_j(Q)] := \mathcal{X}_0.$$

To evaluate the corresponding lower bound $\min_{x \in \mathcal{X}_0} F_0(x, P_t)$ would mean to solve a facial disjunctive program.

Notice that **no convexity assumptions with respect to x were needed.**

Directional derivative

Assume now that problem (1) is **convex** with respect to x . Then directional derivative of optimal value function $\varphi(0)$ can be obtained acc. to Gol'shtein (1970), Theorem 17 applied to Lagrange function

$$L(x, u, t) = F_0(x, t) + \sum_j u_j F_j(x, t)$$

when the set of optimal solutions $\mathcal{X}^*(P) = \mathcal{X}^*(0)$ and the set of Lagrange multipliers $\mathcal{U}^*(P) = \mathcal{U}^*(0)$ are nonempty and compact and all functions F_j are **linear in P** – linearity in the contamination parameter t :

$$\varphi'(0^+) = \min_{x \in \mathcal{X}^*(0)} \max_{u \in \mathcal{U}^*(0)} \frac{\partial}{\partial t} L(x, u, 0) = \min_{x \in \mathcal{X}^*(0)} \max_{u \in \mathcal{U}^*(0)} (L(x, u, Q) - L(x, u, P)). \quad (12)$$

Formula (12) simplifies substantially when $\mathcal{U}^*(0)$ is a singleton. When the constraints do not depend on P we get

$$\varphi'(0^+) = \min_{x \in \mathcal{X}^*(0)} \frac{\partial}{\partial t} F_0(x, 0^+) = \min_{x \in \mathcal{X}^*(0)} (F_0(x, Q) - \varphi(0)). \quad (13)$$

These formulas can be exploited to construct an upper bound. More general cases are treated in e.g. Bonnans-Shapiro

Upper bound

To derive an upper bound for optimal value of the contaminated problem with probability dependent constraints we shall assume that all functions $F_j(x, t)$, $j = 0, \dots, J$, are **linear in t** on interval $[0, 1]$. Denote

$$F(x, P_t) = F(x, t) := \max_j F_j(x, t).$$

For convex $F_j(\bullet, P) \forall j$ the max function $F(\bullet, P)$ is convex and

$$\mathcal{X}(t) = \mathcal{X} \cap \{x : F(x, t) \leq 0\}$$

with one linearly perturbed convex constraint.

1. Assume first that for optimal solution $x^*(0)$ of (1), $F(x^*(0), P) = 0$ and $F(x^*(0), Q) \leq 0$. Then at least one of constraints is active at optimal solution and $x^*(0) \in \mathcal{X}(t) \forall t$:

$$\begin{aligned} F(x^*(0), t) &= \max_j [(1-t)F_j(x^*(0), P) + tF_j(x^*(0), Q)] \\ &\leq (1-t)F(x^*(0), P) + tF(x^*(0), Q) \leq 0. \end{aligned}$$

\rightsquigarrow *trivial global upper bound* $F_0(x^*(0), t) \geq \varphi(t)$; if $F_0(x, P)$ is linear in P

$$\varphi(t) \leq F_0(x^*(0), t) = (1-t)\varphi(0) + tF_0(x^*(0), Q) \forall t \in [0, 1]; \quad (14)$$

Local upper bound via NLP stability results

In **convex case** – analyze optimal value function by 1st order methods:
If $x^*(0)$ is **nongenerate point**, \mathcal{X} in (4) convex polyhedral, the contaminated problem reduces **locally** into problem with **parameter independent set of feasible solutions** e.g. [Robinson] \rightarrow for t small enough optimal value function $\varphi(t)$ is concave and its upper bound equals

$$\varphi(t) \leq \varphi(0) + t\varphi'(0^+) \quad \forall t \in [0, t_0]. \quad (15)$$

Nongenerate point: for $\mathcal{X} = \mathbb{R}^n$ means independence of gradients of active constraints at $x^*(0)$ or nondegeneracy for LP.

If also **strict complementarity** holds true, one faces locally an unconstrained minimization problem. More detailed insight can be obtained by a second order analysis; e.g. if \exists continuous trajectory $[x^*(t), u^*(t)]$ of optimal solutions and Lagrange multipliers of (5) emanating from the unique optimal solution $x^*(0)$ and unique Lagrange multipliers $u^*(0)$ of (1) we get (15) with

$$\varphi'(0^+) = (L(x^*(0), u^*(0), Q) - \varphi(P)). \quad (16)$$

Illustrative example – mean-CVaR models

Consider $S = 53$ equiprobable scenarios of weakly returns ϱ of $N = 9$ assets (9 European stock market indexes: AEX, ATX, BCII, BFX, FCHI, GDAXI, PSI20, IBEX, ISEQ) in period 5.10.2007 - 3.10.2008. The scenarios can be collected in the matrix

$$R = \begin{pmatrix} r^1 \\ r^2 \\ \vdots \\ r^S \end{pmatrix}$$

where $r^s = (r_1^s, r_2^s, \dots, r_N^s)$ is the s -th scenario. We will use x for the vector of portfolio weights and the portfolio possibilities are given by

$$\mathcal{X} = \{x \in \mathbb{R}^N \mid 1'x = 1, x_n \geq 0, n = 1, 2, \dots, N\}$$

that is, the short sales are not allowed. The historical data comes from pre-crisis period. The data are contaminated by a scenario r^{S+1} from 10.10.2008 when all indexes strongly fell down. The additional scenario can be understood as a stress scenario or the worst-case scenario.

Illustrative example – mean-CVaR models

Index	Country	Mean	Max	Min	A.S.
AEX	Netherlands	-0.0098	0.10508	-0.12649	-0.24551
ATX	Austria	-0.01032	0.067022	-0.06982	-0.28503
BCII	Italy	-0.01051	0.047976	-0.06044	-0.19581
BFX	Belgium	-0.00997	0.051099	-0.07386	-0.2253
FCHI	France	-0.00795	0.050254	-0.06292	-0.21704
GDAXI	Germany	-0.00742	0.040619	-0.07568	-0.21151
PSI20	Portugal	-0.00998	0.049866	-0.07404	-0.18116
IBEX	Spain	-0.00625	0.053098	-0.06992	-0.2074
ISEQ	Ireland	-0.01378	0.113174	-0.14689	-0.26767

Table: Descriptive statistics of 9 European stock indexes and the additional scenario

We will apply the contamination bounds to mean-risk models with CVaR as a measure of risk. Two formulations are considered: In the first one, we are searching for a portfolio with minimal CVaR and at least the prescribed expected return. Secondly, we minimize the expected loss of the portfolio under the condition that CVaR is below a given level.

Illustrative example – CVaR minimizing

Mean-CVaR model with CVaR minimization is a special case of the general formulation (1) when $F_0(x, P) = \text{CVaR}(-\varrho'x)$ and $F_1(x, P) = E_P(-\varrho'x) - \mu(P)$; $\mu(P)$ is the maximal allowable expected loss. We choose

$$\mu(P) = -E_P \varrho' \left(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9} \right)' = \frac{1}{53} \sum_{s=1}^{53} -r^s \left(\frac{1}{9}, \frac{1}{9}, \dots, \frac{1}{9} \right)'.$$

It means that the minimal required expected return is equal to the average return of the equally diversified portfolio. The significance level $\alpha = 0.95$ and \mathcal{X} is a fixed convex polyhedral set representing constraints that do not depend on P .

We construct:

- Lower bound (globally for $t \in [0, 1]$):

$$(1 - t) \min\{\varphi(0), \min_{\mathcal{X}(Q)} F_0(x, P)\} + t \min\{\varphi(1), \min_{\mathcal{X}(P)} F_0(x, Q)\}$$

Illustrative example – CVaR minimizing

- Trivial upper bound (globally for $t \in [0, 1]$): Since $x^*(0)$ is a feasible solution of fully contaminated problem, we may use the trivial global bound:

$$F_0(x^*(0), P_t) = \text{CVaR}_\alpha(x^*(0), (1-t)P + tQ)$$

The disadvantage of this trivial bound is the fact, that it would require evaluation of the CVaR for each t . Linearity with respect to t does not hold true, but using concavity of CVaR with respect to t , we may derive an upper estimate for $F_0(x^*(0), t)$:

- Upper estimate of upper bound (globally for $t \in [0, 1]$):

$$\begin{aligned} & \text{CVaR}_\alpha(x^*(0), (1-t)P + tQ) \\ & \leq (1-t)\text{CVaR}_\alpha(x^*(0), P) + t\Phi_\alpha(x^*(0), v^*(x, P), Q), \end{aligned}$$

see [D-P].

Illustrative example – CVaR minimizing

The lower bound is linear, the upper bound is piecewise linear in t and for small values of t it coincides with the estimated upper bound.

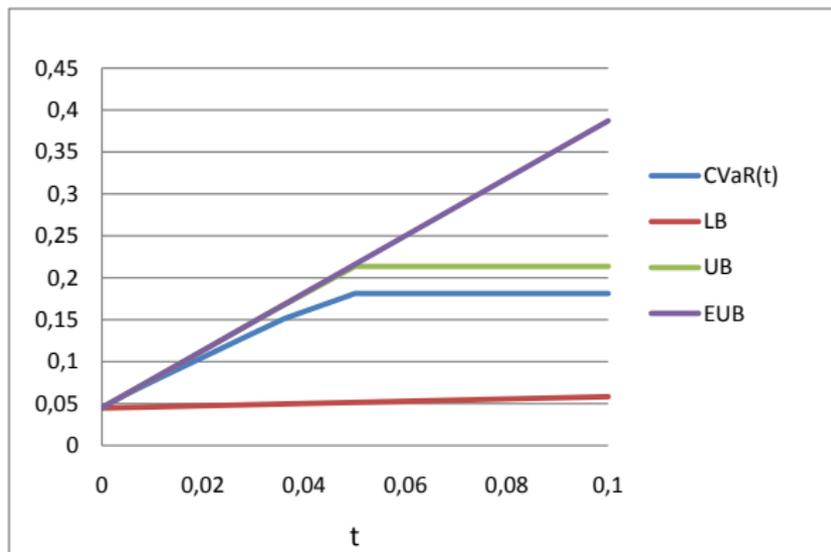


Figure: Comparison of optimal values ($CVaR(t)$) of mean-CVaR models with lower bound (LB), upper bound (UB) and the estimated upper bound (EUB) for the contaminated data.

Illustrative example – Expected loss minimizing

As the second example, consider the mean-CVaR model minimizing the expected loss subject to a constraint on CVaR. This corresponds to (1) with $F_0(x, P) = E_P(-\varrho'x)$ and $F_1(x, P) = \text{CVaR}(-\varrho'x) - c$ where $c = 0.19$ is the maximal accepted level of CVaR. For simplicity, this level does not depend on the probability distribution. Similarly to the previous example, we compute the optimal value $\varphi(t)$ and its lower and upper bound.

- Lower bound (globally for $t \in [0, 1]$):

$$(1 - t) \min\{\varphi(0), \min_{\mathcal{X}(Q)} F_0(x, P)\} + t \min\{\varphi(1), \min_{\mathcal{X}(P)} F_0(x, Q)\}$$

- Upper bound (locally for $t \in [0, t_0]$): In this case $x^*(0) \notin \mathcal{X}(Q)$, hence the trivial upper bound can not be used. Therefore we apply the more general upper bound:

$$\varphi(t) \leq \varphi(0) + t\varphi'(0^+) \forall t \in [0, t_0].$$

that leads to:

$$\varphi(t) \leq (1 - t)\varphi(0) + tF_0(x^*(0), Q) \forall t \in [0, t_0].$$

Illustrative example – Expected loss minimizing

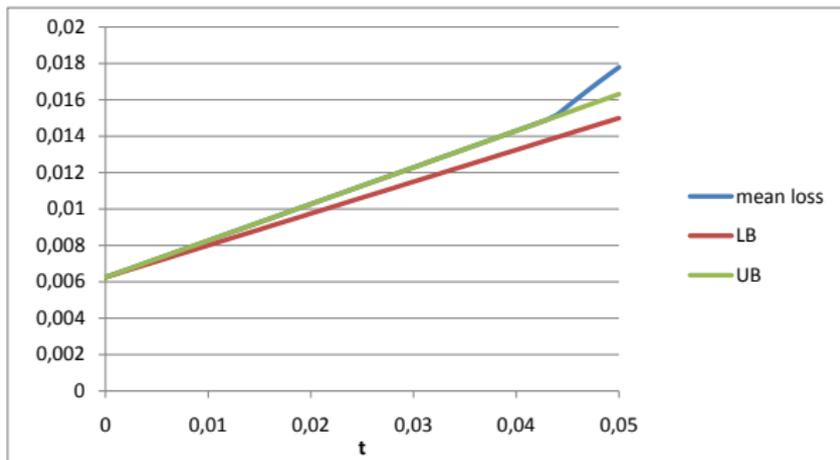


Figure: Comparison of minimal mean loss values with its lower bound (LB) and upper bound (UB) for the contaminated data.

The upper bound coincides with $\varphi(t)$ for $t \leq 0.043$. It illustrates the fact that the local upper bound is meaningful if the probability of the additional scenario is not too large, i.e. no more than the double of probabilities of the original scenarios for our example.

- The contamination technique was extended to construction of bounds for the optimal value function of perturbed stochastic programs whose set of feasible solutions depends on the probability distribution.
- Lower bounds for optimal value function were constructed (objective function depends on P , does not depend on P)
- Upper bounds for optimal value functions derived (trivial global upper bound, general local upper bound)
- Application to mean-CVaR models
- Possible application to CVaR risk-shaping problem, optimization problems with second order stochastic dominance (SSD) constraints, portfolio efficiency testing with respect to SSD criterion, etc.

Dupačová, J., Kopa, M (2012): Robustness in Stochastic Programs with Risk Constraints, *Annals of Operations Research*, 200, 1, 55 – 74.

Dupačová, J., Kopa, M (2014): Robustness of optimal portfolios under risk and stochastic dominance constraints, *European Journal of Operational Research*, 234, 2, 434 – 441.

Main quoted references

- [**Bonnans-Shapiro**] Bonnans, J. S., Shapiro, A. (2000). *Perturbation Analysis of Optimization Problems*. Springer Series in OR, New York.
- resist** Dupačová (1996) Scenario based stochastic programs: Resistance with respect to sample. *Ann. Oper. Res.* 64:21–38
- [**D-P**] Dupačová, Polívka (2007) Stress testing for VaR and CVaR. *Quantitative Finance* 7:411–421
- Kyber** Dupačová(2008) Risk objectives in two-stage stochastic programming models. *Kybernetika* 44:227–242
- [**P-R**] Pflug, Römisch (2007) *Modeling, Measuring and Managing Risk*. World Sci., Singapur
- [**Robinson**] Robinson (1986) Local structure of feasible sets in nonlinear programming Part III: Stability and sensitivity. *Math. Progr. Study* 30:45–66
- [**R-U**] Rockafellar, Uryasev (2002) Conditional value-at-risk for general loss distributions. *J. of Banking and Finance* 26:1443–1471
- [**Romisch**] Römisch (2003) Stability of stochastic programming problems. Chapter 8 in [R-S]
- [**R-S**] Ruszczyński, Shapiro eds. (2003) *Stochastic Programming*. Handbooks in OR & MS